

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日:
2004年7月15日(15.07.2004)

PCT

(10) 国际公布号:
WO 2004/058971 A1

(51) 国际分类号⁷: C12N 15/12, C07K 14/435, C12N 15/63, 15/11, C12Q 1/68, G01N 33/574, 33/577, 33/68

(21) 国际申请号: PCT/CN2003/001109

(22) 国际申请日: 2003年12月24日(24.12.2003)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
02158110.X 2002年12月24日(24.12.2002) CN
03109786.3 2003年4月21日(21.04.2003) CN

(71) 申请人(除美国以外的所有指定国): 北京大学
(PEKING UNIVERSITY) [CN/CN]; 中国北京市海淀区学院路38号北京大学医学部细胞生物学教研室,
Beijing 100083 (CN)。

(72) 发明人;及

(75) 发明人/申请人(仅对美国): 周柔丽(ZHOU, Rouli)
[CN/CN]; 邵根泽(SHAO, Genze) [CN/CN]; 刘歆荣
(LIU, Xinrong) [CN/CN]; 张青云(ZHANG, Qingyun)
[CN/CN]; 芮静安(RUI, Jingan) [CN/CN]; 张页
(ZHANG, Ye) [CN/CN]; 金月英(JIN, Yueying) [CN/
CN]; 林明(LIN, Ming) [CN/CN]; 张莎(ZHANG, Sha)
[CN/CN]; 中国北京市海淀区学院路38号北京大学医学部细胞生物学教研室, Beijing 100083 (CN)。

(74) 代理人: 北京纪凯知识产权代理有限公司(JIEKAI & PARTNERS); 中国北京市西城区宣武门西大街129号金隅大厦602室, Beijing 100031 (CN)。

(81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(84) 指定国(地区): ARIPO专利(BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

本国际公布:

— 包括国际检索报告。

所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: HUMAN CANCER-RELATING GENES, THE PRODUCTS ENCODED THEREBY AND APPLICATIONS THEREOF

(54) 发明名称: 人癌症相关基因及其编码产物与应用

(57) Abstract: The present invention discloses a human cancer-relating gene, *LAPTM4B*, the product it encodes and their applications. The titled human cancer-relating gene provided in this invention is one of the following nucleotide sequences: (1) SEQ. ID. No. 1, SEQ. ID. No. 2, SEQ. ID. No. 3, SEQ. ID. No. 6, or SEQ. ID. No. 8 shown in the sequence listings; (2) Nucleotide sequences that encode the protein sequences of SEQ. ID. No. 4, SEQ. ID. No. 5, or SEQ. ID. No. 7 shown in the sequence listings, and (3) DNA sequences that have more than 90% homology with the DNA sequences specified by the SEQ. ID. No. 1, SEQ. ID. No. 2, SEQ. ID. No. 3, SEQ. ID. No. 6, SEQ. ID. No. 8 shown in the sequence listings, and that encode proteins having the same or similar functions. Based on the present invention, it is possible to develop new anti-cancer pathways and drugs. This would be a project of significant social benefits.

(57) 摘要

本发明公开了人癌症相关基因 *LAPTM4B* 及其编码产物与应用。本发明所提供的人癌症相关基因名称为下列核苷酸序列之一: 1) 序列表中的 SEQ ID No: 1、SEQ ID No: 2、SEQ ID No: 3、SEQ ID No: 6 或 SEQ ID No: 8; 2) 编码序列表中 SEQ ID No: 4、SEQ ID No: 5 或 SEQ ID No: 7 蛋白质序列的多核苷酸; 3) 与序列表中 SEQ ID No: 1、SEQ ID No: 2、SEQ ID No: 3、SEQ ID No: 6 或 SEQ ID No: 8 限定的 DNA 序列具有 90% 以上同源性, 且编码相同或相近功能蛋白质的 DNA 序列。以本发明为基础, 有可能开发出一些新的抗癌途径和抗癌药物。这是一项将产生重大社会效益的工程。

WO 2004/058971 A1

BEST AVAILABLE COPY

人癌症相关基因及其编码产物与应用

技术领域

本发明涉及基因工程和蛋白质工程领域中与癌症相关的基因及其编码产物与应用。

背景技术

癌症是严重危害人民健康的重大疾病，其中肝细胞癌（hepatocellular carcinoma, HCC）尤为凶险。据报道，全世界原发性肝癌每年的新病人数超过 100 万人，其中 70%集中在亚洲。我国肝癌约占全世界肝癌的 40-45%，每年新生肝癌患者 45 万左右，并且呈不断上升趋势，尤其是 20-60 岁的社会中坚人群的发病率显著上升。肝癌不仅发病率高，而且隐匿、进展快、复发率和死亡率高，被称为“癌中之王”。到医院就诊的肝癌患者大多为中或晚期，若不能得到有效治疗自然病程仅 3-6 个月。

阐明癌症，特别是肝癌的发病机制将有助于癌症的预防、诊断和治疗。早期诊断是提高疗效、降低死亡率的关键。目前所用的肝癌诊断标志物 AFP 在约 30%肝癌患者为阴性，而在一些良性肝病患者的血清中又可以出现 AFP 的大幅度升高，造成鉴别诊断的困难。研究发现，癌症的发病与个体的遗传易感性相关。不同遗传背景的个体对环境致癌因子的处理能力存在差异，导致个体患癌风险不同。个体的这种肿瘤遗传易感性不同之本质在于基因的多态型和多样性。

癌症在本质上是细胞的遗传性疾病。与癌症相关的基因虽已发现很多，但癌症发生、发展的机制仍然不甚明了。目前发现的原癌基因按其编码产物在细胞的定位及功能大体可以归为五类，一类是编码生长因子的基因，包括 sis、int-2、hst、fgf-5；第二类是编码生长因子受体的基因，包括 erbB、erbB-2、fms、met、ros 等；第三类是编码细胞质中信号传导分子的基因，包括 abl、src、ras、raf、yes、fgr、fes、lck、mos 等；第四类是编码细胞增殖与凋亡调控分子的基因，包括 bcl-1、bcl-2 等；第五类是编码细胞核里同 DNA 相结合的蛋白质（转录因子）的基因，例如 myc、myb、fos、jun、B-lym、ski、ets、rel 等基因。研究证实，和 HCC 的发生密切相关的主要有 ras、src、myc、met、p53 等基因。

发明公开

本发明的目的是提供一个新的人癌症相关基因及其编码产物。

本发明所提供的人癌症相关基因名称为 *LAPTM4B*，是下列核苷酸序列之一：

1、人癌症相关基因，是下列核苷酸序列之一：

1) 序列表中的 SEQ ID No: 1、SEQ ID No: 2、SEQ ID No: 3 或 SEQ ID No: 6;

2) 编码序列表中 SEQ ID No: 4、SEQ ID No: 5 或 SEQ ID No: 7 蛋白质序列的多核苷酸;

5 3) 与序列表中 SEQ ID No: 1、SEQ ID No: 2、SEQ ID No: 3 或 SEQ ID No: 6 限定的 DNA 序列具有 90% 以上同源性, 且编码相同或相近功能蛋白质的 DNA 序列。

所述序列表中的 SEQ ID No: 1 由 954 个碱基组成, 是一个完整的读码框, SEQ ID No: 1 具有两个起始位点, 一个是自 5' 端 1—3 位碱基, 另一个是自
10 5' 端 274—276 位碱基。序列 SEQ ID No: 1 的完整 cDNA 有两个, 具有不同的加尾信号, 当 SEQ ID No: 1 的 5' 端向外延伸 85 个碱基, 3' 端向外延伸 401 个碱基, 即得到序列表中的 SEQ ID No: 2, 该基因由 1440 个碱基组成; 当 SEQ ID No: 1 的 5' 端向外延伸 85 个碱基, 3' 端向外延伸 1130 个碱基, 即得到序列表中的 SEQ ID No: 3, 该基因由 2169 个碱基组成。基因 *LAPTM4B*
15 定位于染色体 8q22.1。

序列表中的 SEQ ID No: 6 是 SEQ ID No: 1 的等位基因, 由 2264 个碱基组成, 其读码框为自 5' 端第 17 到 1129 位碱基。其中含有两个拷贝的 19bp DNA 片段, 19bp DNA 片段的序列为 gcttgg agctccagca gct。这两个拷贝的 19bp DNA 片段在 SEQ ID No: 1 等位基因 nt 124—nt 161。

20 人类癌症相关蛋白 LAPTM4B, 是具有序列表中序列 4 或/和序列 5 或/和序列 7 氨基酸序列的蛋白质, 或者是将序列 4 或/和序列 5 或/和序列 7 的氨基酸序列经过一个或几个氨基酸残基的取代、缺失或添加且具有与序列 4 或/和序列 5 或/和序列 7 的氨基酸序列相同或相近活性的由 4 或/和序列 5 或/和序列 7 衍生的蛋白质。

25 序列表中的序列 4 由 317 个氨基酸残基组成, 由 SEQ ID No: 1 的全序列编码, 分子量 35kDa, 推定等电点为 9.05; 序列表中的序列 5 由 226 个氨基酸残基组成, 由 SEQ ID No: 1 中第 274—954 位碱基编码, 分子量为 24kDa, 推定等电点为 4.65。序列表中的序列 7 是由 370 个氨基酸残基组成的蛋白质。

LAPTM4B 基因在 16 种正常组织具有有不同程度的广泛表达, 其转录水平
30 以睾丸、心肌及骨骼肌的表达最高, 在卵巢、肾和胰表达中等, 在肝、脾、小肠、大肠和胸腺表达较低, 在肺和外周血细胞表达最低; 在 8 种胎儿组织中以心、骨骼肌及肾的表达水平最高, 胎肝的表达水平略高于成年肝。然而, 在某些癌组织中的表达明显上调, 例如, Northern Blot 分析证明在 87.3% (48/55) 人肝癌组织中的转录水平显著高于胎肝及正常肝 (图 1-A)。原位杂

交(图 2-A)、免疫组化(图 2-B)及免疫细胞化学(图 2-C)也证明其在肝癌组织中特异性高表达,而在配对的非癌肝组织表达较低(图 2-A 及图 2-B)。在所检测的 5 种肝癌细胞系中除 HLE 之外全部以高水平表达,包括 SMMC-7721, QGY-7701, BEL7402 和 HG116(图 1-B 和图 2-C)。重要的是,在肝癌组织和肝癌细胞系高度过表达的蛋白质产物主要是 SEQ ID No: 4 LPTM4B-35;而 SEQ ID No: 5 LPTM4B-24 在肝癌只有轻度的表达上调,以致 LPTM4B-35 与 LPTM4B-24 蛋白的比值在肝癌组织显著升高(图 2-B)。在某些配对非癌肝组织虽然可见 LPTM4B-35 和 LPTM4B-24 的轻度升高但二者的比值与正常肝脏无异(表 1 所示)。这或许是肝癌之癌前阶段的特点。再者, *LPTM4B* 基因的 mRNA 和蛋白质表达水平均与肝癌组织的分化状态呈负相关:低分化者表达很高,而高分化者表达较低(图 1-C)。为了确定 *LPTM4B* 基因与其它癌症的关系,经 Western Blot 及免疫组化鉴定, LPTM4B-35 蛋白在某些上皮来源的癌组织和细胞系,如胃癌、乳腺癌及高转移人肺癌和前列腺癌等也见表达上调(图 11);而且在同源的人肺癌及前列腺癌细胞系还证明其在高转移潜能的细胞中表达高度上调。而在黑色素瘤,无论原位癌或转移癌细胞均未见明显表达。此外,在大鼠及小鼠的成年肝组织中虽有低水平表达,而小鼠腹水型肝癌的表达并无显著上调;在增殖状态正常的大鼠再生肝中表达亦无上调。

表 1 肝癌、配对非癌肝组织及正常肝组织 LPTM4B-35 和 LPTM4B-24 的表达比值

	HCC	PNL	NL
LPTM4B-35	13.32±1.98	4.58±1.31	2.78±0.11
LPTM4B-24	3.59±1.78	1.76±1.24	1.00±0.02
LPTM4B-35/ LPTM4B-24 (ratio)	3.71	2.60	2.78

$P<0.01$ vs. PNL and NL

SEQ ID No: 4、SEQ ID No: 5 及 SEQ ID No: 7 LPTM4B 蛋白均具有四段穿膜序列,一个 N 糖基化位点和典型的溶酶体及胞内体定位信号,皆属于四次穿膜的蛋白质超家族,但分别具有不等的磷酸化位点。实验表明,SEQ ID No: 4 LPTM4B-35 能与质膜中的整合素 $\alpha 6 \beta 1$ (细胞外基质层粘连蛋白的单特异性受体)及表皮生长因子受体 EGFR 形成复合物(图 14-A, B 及 C)并在细胞中共同定位于质膜,可能具有将来自于细胞外基质及生长因子两方面的增殖信号在质膜耦联在一起的功能。这将进一步阐明正常真核细胞生长的定着依赖性(即除需要来自于生长因子的刺激信号之外,还需要来自于一定的细胞外基质的刺激信号才能启动细胞增殖)的分子机制,在阐明细胞增殖的调控机制上具有突破性意义。实验证明, LPTM4B 蛋白 C 端胞质区的酪氨酸

残基 (Tyr₂₈₅) 可发生磷酸化 (图 15-A)。当细胞在层粘连蛋白基质上黏附时其磷酸化水平呈峰形升高 (图 15-A), 并可被 LPTM4B-EC2-pAb 抗体几近完全抑制 (图 15-B), 而无关抗体并无此作用 (图 15-C)。Tyr₂₈₅ 磷酸化以后便形成与胞内信号分子 SH2 结构域相结合的位点; 同时, LPTM4B 的 N 端及 C 端序列具有 Pro 富集区及典型的 SH3 结构域的结合位点。这些资料指明 SEQ ID No: 4 LPTM4B-35 蛋白可能是一个在信号转导中具有重要作用的停靠蛋白或细胞质膜特定微域 (microdomain) 的组织者, 可以募集细胞内、外相关的信号分子以完成与细胞增殖、分化及凋亡相关的信号转导。实验证明, 以 SEQ ID No: 4 的 cDNA 转染小鼠 NIH3T3 细胞系及 HLE 人肝癌细胞系, 可获得稳定转染、过表达 LPTM4B-35 的 NIH3T3-AE 和 HLE-AE 细胞系。通过生长曲线 (图 4)、³H-TdR 参入 (图 5) 及细胞周期 S 相细胞数 (图 6) 均证明细胞增殖速度显著增加; 而且, 转染细胞的增殖对血清中生长因子的依赖性降低, 并可在软胶中大量生成集落; NIH3T3-AE 细胞接种于 NIH 小鼠还可生成中等恶性度的纤维肉瘤 (图 7)。指明 LPTM4B-35 的过表达引起了细胞的增殖失控。再者, HLE-AE 细胞的迁移能力增强; 侵袭人工基底膜 (Matrigel) 的能力也显著增强, 指明 LPTM4B-35 的过表达促进细胞恶性表型的发展。反之, 以 SEQ ID No: 5 的 cDNA (缺少 LPTM4B-35 N 端 91 氨基酸的编码序列) 转染鼠 BHK、NIH3T3 细胞系及人 HLE 肝癌细胞系, 则不能长期存活。这些结果表明: SEQ ID No: 4 LPTM4B-35 蛋白 N 端 91 个氨基酸残基对于其调节细胞增殖的功能至关重要; LPTM4B-35 蛋白与 LPTM4B-24 蛋白具有相互制约的不同功能: LPTM4B-35 的过表达促进了细胞的恶性转化, 而 LPTM4B-24 的过表达促进细胞的死亡; 二者表达的平衡与调控对于恶性肿瘤的发生、发展至关重要; *LPTM4B* 基因可能属于原癌基因。在治疗上, 阻遏 SEQ ID No: 4 LPTM4B-35 及增强 SEQ ID No: 5 LPTM4B-24 的表达可能会抑制肝癌的生长, 逆转其恶性表型或延缓肝癌的发展。此外, LPTM4B-35 的过表达也促进 cyclinD1 (图 13-A) 和 cyclinE (图 13-B) 等增殖调节蛋白表达的上调以及 c-Myc (图 13-C)、c-Jun (图 13-D)、c-Fos (图 13-E) 等原癌基因表达的上调。

针对 SEQ ID No: 4 LPTM4B-35 蛋白各表位的单克隆及多克隆抗体, 例如已经制得的针对 SEQ ID No: 4 LPTM4B-35 的第二胞外区的多抗 LPTM4B-EC2₂₃₂₋₂₄₁-pAb 及针对 SEQ ID No: 4 LPTM4B-35 N 端序列的多抗 (LPTM4B-N₁₋₉₉-pAb 和 LPTM4B-N₂₈₋₃₇-pAb) 以及各种针对 LPTM4B 的单抗均对于研究 LPTM4B-35 和 LPTM4B-24 的功能以及肿瘤的诊断、治疗具有重要作用 (图 2, 3, 8, 11, 12, 14, 15)。例如, LPTM4B-EC2₂₃₂₋₂₄₁-pAb, LPTM4B-N₁₋₉₉-pAb 多抗和 LPTM4B-N₁₋₉₉-mAb 单抗可用于鉴定 LPTM4B 蛋白的

表达、胞内定位、分离纯化、蛋白质相互作用,检测血液中 LPTM4B 抗原及抗体(图 8); LPTM4B-EC2₂₃₂₋₂₄₁-pAb 还可抑制肿瘤细胞的增殖(图 12)和 LPTM4B 蛋白的 Tyr₂₈₅ 的磷酸化(图 15-B)、胞内信号分子 FAK 和 MAPK 的磷酸化与活化(图 16-B)等。因此,针对 SEQ ID No: 4 LPTM4B-35 蛋白各表位的所有单克隆及多克隆抗体均属于本发明的保护范围。

SEQ ID No: 8 是 *LPTM4B* 基因的启动子序列。为了研究 *LPTM4B* 基因的表达调控,克隆了 *LPTM4B* 基因的启动子及其上游序列 SEQ ID No: 8。在 *LPTM4B* 基因启动子区域无典型的 CCAAT (TTGCGCAAT)、TATA 盒;在 *LPTM4B* 启动子上游区域存在多种转录因子结合位点,如 CREBP1/c-Jun, CEBP, PAX2/5/8, GATA, STAT, c-Ets-1, E₂F, LYF-1, and c/v-Myb (图 17A 部分)。这些转录因子可能分别在不同组织细胞调控 *LPTM4B* 的表达。在一些肿瘤中,这些转录因子的表达失控可能导致 *LPTM4B* 的表达失衡。另外,在 *LPTM4B* 启动子上游区域有两个高度同源的重复序列,它是否与 *LPTM4B* 表达调节有关值得研究。通过构建包含不同长度的 *LPTM4B* 启动子上游序列—启动子—5' UTR-35bp 编码区—荧光素酶报告基因载体系列,并将各载体分别转染人肝癌 BEL7402 细胞和 HLE 细胞。如图 17 所示,各种载体转染的细胞都有不同强度荧光素酶活性,表明这些片段都有转录活性,其中最小片段为转录起始位点上游约 38 bp 的一段 DNA (*pGL3-PF4*),它具有基础启动活性,为 *LPTM4B* 基因核心启动子。*pGL3-PF1* 转染体在 BEL7402 的活性为参照启动子 SV40 的 20%,而在 HLE 活性较低,仅为 SV40 的 6%,二者相差约 3 倍。这个数值部分反映了 *LPTM4B* 启动子在这两种细胞中的天然活性,与 Northern blot 结果所反映的 BEL-7402 和 HLE 细胞系分别高和低表达的 mRNA 水平相吻合。另外, *pGL3-PF4* 转染体在这两种细胞中的活性相差悬殊,在 BEL7402 的活性比 HLE 高近 7 倍。以上结果显示, *LPTM4B* 基因在 BEL7402 和 HLE 细胞中的表达和转录调节机制是不同的。

为了确定 *LPTM4B* 的不同基因型与肝癌易感性的关系,本发明对基因组 DNA 进行了 *LPTM4B* 基因分型。将原来克隆的人肝癌相关基因 *LPTM4B* 命名为 *LPTM4B*1*。通过 PCR 克隆到另一等位基因 *LPTM4B*2*,即 SEQ ID No: 6。如图 9 所示,等位基因 *1 与 *2 的区别在于第一外显子 5' UTR 内的一个 19 bp 序列, *1 等位基因只有一个这样的序列 (nt 124 ~ 142);而 *2 含有 2 个,且紧密串联 (124-142dup,以转录起始位点 TSS 处 G 为 +1 计数标准)。由于 19-bp 序列的插入,其对应于 *1 等位基因 5' UTR 中的终止密码因三联移位而失效,导致蛋白质编码框可能在 N 端向上游延长了 53 个氨基酸,因此 SEQ ID No: 6 编码的蛋白质应由 370 个氨基酸残基组成 (SEQ ID No: 7)。在人群中检测到

的 *LAPTM4B* 的基因型分别为 *1/*1、*1/*2 和 *2/*2 (图 10)。研究发现 *LAPTM4B* 基因型 *2/*2 的个体患肝癌的风险比非 *2/*2 型个体高 2.89 倍 (表 2)，而食管癌患者的 *LAPTM4B* 基因分型与正常人群并无差异 (表 3)。表明 *LAPTM4B* *2/*2 基因型特异的与肝癌的易感性相关。因此本发明提供的 *LAPTM4B* 的等位基因 *LAPTM4B**2 可以作为筛查肝癌易感人群和高危人群之目标物的应用，特别是基因型为 *LAPTM4B**2/*2 者作为筛查肝癌易感和高危人群之靶标物的准确性更大。因此，*LAPTM4B* 基因的 *1/*1， *1/*2 及 *2/*2 分型以及由 *LAPTM4B**2 编码的蛋白质或其抗体或者自人基因组中扩增捕获 *LAPTM4B* 的物质为活性成分的试剂可用于制备筛查肝癌高危和易感人群的试剂。

10 含有 SEQ ID No: 1、2、3、6、8 所述序列的表达载体，含有 SEQ ID No: 1、2、3、6、8 序列的转染细胞系及扩增 SEQ ID No: 1、2、3、6、8 的引物也均属于本发明的保护范围。

附图说明

图 1-A 为 Northern Blot 分析图谱，显示本发明基因在人正常肝、胎肝及肝癌组织中的表达 (转录)。

图 1-B 为 Northern Blot 分析图谱，显示本发明基因在人肝癌细胞系中的表达 (转录)。

图 1-C 为显示本发明基因在人肝癌组织中的表达水平与组织分化程度相关性的散点图。

20 图 2-A 为肝癌的原位杂交图，肝癌癌巢中 *LAPTM4B* mRNA 显示强阳性。

图 2-B 为肝癌的免疫组化图，肝癌癌巢中 *LAPTM4B* 蛋白显示强阳性。

图 2-C 为免疫细胞化学图，显示 *LAPTM4B* 蛋白在转染细胞的存在。

图 3 为 Western Blot 分析图，显示本发明基因编码的 *LAPTM4B*-35 和 *LAPTM4B*-24 蛋白在正常肝、肝癌及配对非癌肝组织中的表达谱。

25 图 4 为显示本发明 cDNA 转染细胞的增殖加速的生长曲线。

图 5 为显示本发明 cDNA 转染细胞的 DNA 合成增加的柱图。

图 6 为显示本发明 cDNA 转染细胞的 S 期细胞数增加的饼图 (流式细胞分析结果)。

图 7 为本发明 cDNA 转染的细胞对小鼠的致癌作用。

30 图 8 为肝癌患者血清中的本发明抗原水平直方图。

图 9 为本发明基因 *LAPTM4B* 等位基因的部分片段。

图 10 为本发明基因 *LAPTM4B* 在人群中的基因分型。

图 11 为不同上皮源癌组织的免疫组化图。

图 12 为显示抗体 *LAPTM4B*-EC2-pAb 对肝癌细胞增殖之抑制作用的柱形图。

图 13-A, B, C, D, E 分别为显示本发明 cDNA 转染细胞的 cyclin D1, cyclinE, c-Myc, c-Fos, c-Jun 表达上调的 Western Blot 图。

图 14-A, B, C 为免疫共沉淀分析图谱, 分别显示本发明基因产物与 $\alpha 6 \beta 1$ 整合素以及生长因子受体的相互作用。

5 图 15-A, B, C 为免疫沉淀分析图谱, 显示 LAPTM4B 蛋白的 Tyr 磷酸化及 LAPTM4B-EC2-pAb 对磷酸化的抑制作用。

图 16-A, B 为免疫共沉淀分析图谱, 分别显示 LAPTM4B 参与 FAK-MAPK 信号传导途径。

图 17 为 LAPTM4B 启动子活性图。

10 实施发明的最佳方式

患者及其正常对照组的来源

所有 57 例肝癌患者年龄在 35—70 岁之间, 平均年龄为 54 ± 6.0 岁, 男性患者 50 名, 女性 7 名。实验组织来自外科治疗切除标本或血液。对照组包括 206 名年龄相配的无症状和临床检验没有患肝癌人群的血液及 209 名新生儿出生时的脐静脉血作为样本。

所有 109 名食道癌患者年龄在 30—70 岁之间, 平均年龄为 55 ± 5.4 岁, 男/女比例为 76/33。实验组织来自外科治疗切除标本。正常对照组 S 为 116 名无症状和临床检验没有患食道癌人群的血液。所有样品被用来分离提取基因组 DNA。

20 统计方法

采用 χ^2 检验和单因素 ANOVA 方差分析处理数据

实施例 1、Northern Blot 分析 LAPTM4B 在增殖/分化状态不同的四种肝组织中的表达

25 选择增殖/分化状态不同的四种肝组织, 即正常成年肝 (极少增殖, 高度分化)、胎肝 (旺盛增殖, 低度分化)、肝癌 (失控增殖, 异常分化) 及配对非癌肝组织 (活跃增殖的癌前肝细胞) 进行 Northern Blot 分析, 从 5 例手术治疗切除的正常成年肝组织, 5 例流产胎儿的胎肝组织, 55 例手术治疗切除的肝癌组织和 55 例配对非癌肝组织的新鲜标本提取 RNA, 经电泳分离、转移至尼龙膜, 以 Dig 标记的探针进行杂交, 于 68°C 洗膜, 按说明书显示杂交信号。结果如图 1 所示, 其中泳道 1 是胎肝, 泳道 2 是正常成年肝, 泳道 3、5、7、9 是肝癌, 泳道 4、6、8、10 是配对非癌肝组织。结果表明, LAPTM4B 在各种肝组织的表达水平依次为: 肝癌 > 配对非癌肝组织和胎肝 > 正常成年肝。

实施例 2、*LAPTM4B* 基因及其启动子与等位基因的克隆

一、*LAPTM4B* 基因的克隆

选择增殖/分化状态不同的四种人类肝组织，即正常成年肝 (NL)、胎肝 (FL)、肝癌 (HCC) 及配对非癌肝组织 (PNL)，通过荧光差异显示技术得到了
5 一个未知基因的 cDNA 片段 (LC27)，以 LC27 片断 (426bp) 为基础向 5' 方向进行 EST 同源序列拼接，并通过 RACE 及 高温 RT-PCR 实验得到基因的全长 cDNA 序列，即为 SEQ ID No: 2 和 3。

二、*LAPTM4B* 启动子的克隆

应用生物信息学方法对基因组 *LAPTM4B* 基因第一外显子 5' 端上游序列
10 进行分析，设计引物 F1 和 R1，以人基因组 DNA 为模板，用超高保真 DNA 聚合酶 Pfx 扩增得到了 *LAPTM4B* 基因的启动子及其上游序列。经 *Xho* I 和 *Hind* III 酶切后连接到 pGL3-Basic 载体，得到 *pGL3-PF1*，测序鉴定 (测序结果见图 16 中的 a 部分)。

如图 17 -A 部分所示，*LAPTM4B* 基因启动子区域无典型的 CCAAT
15 (TTGCGCAAT)、TATA 盒。在 *LAPTM4B* 启动子上游区域还存在多种转录因子结合位点，如 CREBP1/c-Jun, CEBP, PAX2/5/8, GATA, STAT, c-Ets-1, E2F, LYF-1, and c/v-Myb。它们可能与 *LAPTM4B* 的表达调控有关。在肝癌中，这些转录因子的表达失控可能导致 *LAPTM4B* 的表达失衡。另外，在 *LAPTM4B* 启动子上游区域有两个高度同源的重复序列，它是否与 *LAPTM4B* 表达调节有关值得研究。

三、*LAPTM4B* 等位基因的克隆和测序

(1) DNA 的分离

根据标准的苯酚-氯仿法，从正常人、肝癌或食道癌患者的血液淋巴细胞或外科治疗切除组织标本中提取基因组 DNA。

(2) 等位基因的克隆和测序

与启动子序列的克隆相同，根据 *LAPTM4B* 基因 SEQ ID No: 3，设计合成的两个引物，F₁: 5' GCGCTCGAGGCTCCAGGTG GAAGAGTGTGC 3' (5' 末端引入
25 *Xho*I 酶切位点，即划线部分)；R₁: 5' GCGAAGCTT GGACTTGGCCATGTGACCCG 3' (5' 末端引入 *Hind*III 酶切位点，即划线部分)，通过 PCR 法从人基因组 DNA 中克隆 *LAPTM4B* 第一外显子的启动子及其前续部分的序列。构建不同人基因组 DNA 来源的 *pGL3-PF1* 载体进行测序，筛选 *LAPTM4B* 等位基因。将原来的 *LAPTM4B*
30 序列命名为 *LAPTM4B*1* 序列。 *LAPTM4B*2* 的序列为序列表中 SEQ ID No: 6。图 9-A 为 *LAPTM4B* 启动子和第一外显子示意图。长方形框表示第一外显子，黑色部分为编码区，白色部分为非编码区，灰色部分为 19bp 的 DNA 序列，横线表示启动子，F₁、F₂、R₁、R₂ 分别为四条引物所在位置。该序列以起始密码子 ATG

中的A定为+1。图9(B)为*LAPTM4B*等位基因的部分序列和色谱图。用下划线标记的是19bp的DNA序列。结果表明, *LAPTM4B*1*中含有一个拷贝的19bp DNA序列, *LAPTM4B*2*含有两个拷贝的19bp DNA序列, 并且它们衔接在*LAPTM4B*1*第一外显子的5' 非编码区(nt -33--15)。

- 5 测序结果表明 *LAPTM4B*2* 和 *LAPTM4B*1* 的启动子相同, 并没有发现 *LAPTM4B* 等位基因*1和*2的启动子序列存在差别。

(3) *LAPTM4B* 基因分型

- 设计引物 E₂ (5' GCCGACTAGGGGACTGGCGGA 3') 和 R₂ (5' CGAGAGCTCCGAGCTTCTGCC 3'), 以正常人、肝癌、食管癌患者基因组 DNA 为模板, 扩增 *LAPTM4B* 第一外显子部分序列。PCR 条件如下: 96°C 预变性 5 min; 94°C 30 s, 68°C 30 s, 72°C 1 min, 35 循环; 72°C 延伸 5 min。2%琼脂糖凝胶电泳分析。图 10 表明在人群中存在 *LAPTM4B* 基因的*1/*1, *1/*2*及 2/*2 三种类型。

- 15 实施例 3、报告质粒的构建和启动子活性的分析

- 构建包含 *LAPTM4B* 不同长度启动子上游序列—启动子—5' UTR—35bp 编码区—荧光酶报告基因载体系列, 即将 *LAPTM4B* 基因的启动子及其上游序列, 经 *Xho* I 和 *Hind* III 酶切后连接到 pGL3-Basic 载体, 得到 *pGL3-PF1*, 测序鉴定。以 *pGL3-PF1* 为模板, 用引物 F2、F3、F4 分别与 R1 扩增不同长度启动子片段—荧光素酶载体 *pGL3-PF2*、*pGL3-PF3*、*pGL3-PF4*, 测序鉴定。

引物序列如下:

- F1: 5' GCGCTCGAG GCTCCAGGTGGA AGAGTGTGC 3' (nt -1341- -1321)
 F2: 5' GCGCTCGAG TAA AAACGCTGTGCCAGGCGT 3' (nt -881- -861)
 F3: 5' CCGCTCGAG TACCGGAAGCACAGCGAGGAT 3' (nt -558- -538)
 25 F4: 5' GCGCTCGAG AGTAGAAGGGAAGAAAATCGC 3' (nt -38- -18)
 R1: 5' GCGAAGCTT GGAAGCTTGGCCATGTGACCCG 3' (nt 172-191)

- 将上述各载体分别转染 BEL7402 细胞以及 HLE 细胞, 测定其启动子活性。如图 17 的 b 部分所示, 各种载体转染的细胞都有不同强度荧光酶活性。*pGL3-PF3* 在这两种细胞中的活性接近, 都为 SV40 启动子 (*pGL3-Promoter*) 活性的 27%。但它与 *pGL3-PF4* 的活性相比时, 在 BEL7402 细胞几乎无差别, 在 HLE 细胞中前者显著高于后者 (相差 7 倍)。从图 17 的 a 部分可以看到, 在 *pGL-PF3* 上 (-41~-558) 有众多潜在的转录因子结合位点, 它们之一或多个, 尤其是 c-Ets-1, 可能在 HLE 细胞中发挥了调节作用, 导致 HLE 细胞 *pGL3-PF3* 和 *pGL3-PF4* 转染体荧光酶活性的显著差异。在 BEL7402 和 HLE 细

胞中, *pGL3-PF3* 活性比 *pGL3-PF1*、*pGL3-PF2* 都要高, 提示某种(些)负性调控因子的存在, 它(们)与启动子上游靶序列(-558 上游)的结合导致了 *LAPTM4B* 基因的下调表达; 尤其是在 HLE 细胞, 这种抑制作用表现得更为明显, 表明在 HLE 细胞中可能存在某种因子强烈抑制 *LAPTM4B* 的表达。图 1-B 所示 Northern Blot 结果也与此相符, 即 *LAPTM4B* 在 HLE 细胞中 *LAPTM4B* 表达很低。载体 *pGL3-PF2* 包含两个 DNA 重复片段(-41~-328, -574~-859), 比 *pGL3-PF3* 多出一个 DNA 重复片段(-574~-859), 而在两种细胞中 *pGL3-PF3* 活性比 *pGL3-PF2* 都要高。这表明, 这两个重复序列对基因转录有负性调控的作用。这两段序列上有多个潜在的转录因子结合序列, 它们给每个负性调控因子提供了两个结合位点, 而很多转录调节因子往往形成二聚体, 同时与两个靶序列结合位点相结合才能发挥功能, 如果象 *pGL3-PF3* 这样只提供一个结合位点, 就不能发挥作用。因此, *pGL3-PF3* 转染体由于去抑制, 其活性比其它载体转染体要高。

实施例4、LAPTM4B蛋白表达的Western Blot分析

将组织置于冰上, 用剪刀将其剪成碎片, 挑取约 0.1g 湿重组织放入手动匀浆器中, 每管加入 1ml 裂解液, 于冰浴中充分匀浆。将裂解液转移至离心管中, 12 000g 4℃离心 10min, 以去除未完全裂解的组织碎片。或将培养瓶中的细胞以 0.25%胰酶消化收获, PBS 洗涤两次, 裂解细胞, 500g 离心 3 min 收集上清。12% SDS-PAGE 电泳分离, 转膜。于 4℃封闭液(TBS, 0.05% Tween 20, 5% 脱脂奶粉)中封闭过夜。然后用兔 LAPTM4B-EC2₂₃₂₋₂₄₁-pAb 多克隆抗体(1:500 稀释)或小鼠 Anti-FLAG M₂单抗(Sigma, 1:750 稀释)室温孵育 2 h。TBS 洗涤 3 次后, 分别用过氧化物酶偶联的二抗即羊抗兔或羊抗鼠 IgG (1:3000 稀释)室温孵育 2 h。洗涤缓冲液(TBS, pH 8.0, 0.05% Tween 20)洗 3 次, 最后一次用不含 Tween 20 的洗涤缓冲液洗涤 1 次。ECL (Santa Cruz)显色曝光(依厂家说明进行)。当同一张膜需要杂交两个抗体时, 可用先用 TBS 漂洗经 ECL 曝光后的膜, 再用 30 ml TBS (含 2% SDS, 210 μl β-巯基乙醇)室温洗膜 30 min, TBS 洗膜 30 min 即可洗去原先抗体及信号, 用于第二次杂交。结果如图 3 所示, 表明 LAPTM4B-35 在肝癌组织及细胞系过表达。

实施例 5、全长 cDNA 转染实验证明本发明基因对细胞增殖及肿瘤细胞恶性表型的调控作用

通过PCR, 以 *pGEMT-E₂E₇* 质粒为模板, 分别以引物A 或 B 与 E, 以超高保真 Pfx DNA聚合酶扩增 *LAPTM4B* 基因全长或部分读码框。在引物 A 和 B 5' 端引入 *Ban*HI 酶切位点(GGATCC)以及核糖体结合位点序列(GCCACC), 在

引物 E 中引入 *EcoRI* 酶切位点 (GAATTC)。扩增产物 AE 和 BE 经 *Bam*HI、*EcoRI* 酶切、纯化后, 连接与 *pcDNA3.0* 载体相应的多克隆位点, 常规转化 DH5 大肠杆菌, 筛选阳性克隆, 测序鉴定。所构建载体分别命名为 *pcDNA3/AE* 和 *-BE*。其中 *pcDNA3/AE* 包含基因全长 ORF; *pcDNA3/BE* 只含 ORF 从第二个
5 ATG 到 TAA 的序列, 与 *pcDNA3/AE* 相比, 它所编码的蛋白缺少 N 端的 91 个氨基酸。

以 *pcDNA3/AE* 和 *-BE* 分别对低表达 *LAPTM4B* 的鼠 BHK、NIH3T3 细胞系和人肝癌 HLE 细胞系进行转染, 筛选稳定高效表达 *LAPTM4B* 的克隆, 通过酸性磷酸酶法测定活细胞数, 绘制细胞生长曲线; 通过流式细胞术检测细胞周期;
10 以 Western Blot 检测 cyclin D1 和 cyclin E 细胞周期蛋白和 c-Myc, c-Fos, c-Jun 等原癌基因(调节细胞增殖的转录因子)的表达水平。结果表明, 以 *LAPTM4B-AE* 表达质粒转染的细胞, 细胞增殖加速(图 4, 5, 6); 细胞周期蛋白 cyclin D1 和 cyclin E 以及 c-Myc, c-Fos, c-Jun 原癌基因的表达明显增强(图 13-A, B, C, D, E); *LAPTM4B-35* 过表达细胞对血清的依赖性显著降低
15 (在 1%FCS 中 HLE-AE 细胞的增殖照样进行, 而 HLE 和 HLE-MOCK 细胞则几乎不能增殖); 同时, HLE-AE 细胞生长的定着依赖性也显著减弱, 可在琼脂软胶中大量形成大的集落, 表明该基因参与细胞增殖的调控, 其过表达(活化)与细胞的增殖失控有关。再者, HLE-AE 细胞的迁移能力增强(迁移穿过膜孔的细胞数从对照的 1216.5 ± 403.8 增加为 4082.5 ± 748.8); 侵袭人工基底膜
20 (Matrigel) 的能力也显著增强(从对照的 25 ± 12.73 增加为 1325 ± 424.26), 指明 *LAPTM4B-35* 的过表达促进细胞恶性表型的发展。反之, 以 *LAPTM4B-BE* 表达质粒转染的 BHK-BE、NIH3T3-BE 和 HLE-BE 细胞则不能形成克隆, 于 3 周内全部死亡。证明 *LAPTM4B-24* 具有与 *LAPTM4B-35* 相拮抗的作用。

实施例 6、cDNA 转染的细胞对小鼠的致瘤作用

25 随机取 6 周龄 NIH 雄性小鼠分为三组: 生理盐水注射的对照组, 空载质粒转染的 MOCK 细胞接种的对照组和含全长 cDNA 质粒转染的 NIH3T3 细胞接种的实验组。每只小鼠于右腋皮下接种 2×10^6 细胞。每组 4-6 只。21 天后断颈处死, 进行解剖。结果如图 7 所示, 可以看到实验组有两只小鼠形成了明显的中等恶性度的纤维肉瘤(A、B), 另两只的接种部位为淋巴组织(C、D)。
30 而两个对照组的 12 只小鼠至接种 86 天后均无肿瘤生成。

实施例 4, 5, 6 的结果表明, *LAPTM4B* 可能是一个新的原癌基因。

实施例 7、以 ELISA 初步分析肝癌患者血清中的 *LAPTM4B* 抗原水平

将 96 孔板每孔加入不同稀释倍数的 HCC 患者血清或正常人血清包被, 4℃ 过夜。各孔用含 0.5% Tween-20 的 PBS 液漂洗后, 加入 2% BSA 液于室温封闭

1hr。加入不同稀释倍数的LAPTM4B-EC2-pAb抗体，置于室温2hr，用PBS液漂洗后加入辣根过氧化物酶标记的羊抗兔抗体（1：1000倍稀释）置于室温2hr。经PBS漂洗，加入1 μ g/ml邻苯二胺显色10-15分钟后，加入H₂SO₄终止反应，用酶标仪于490nm 波长下读数，检测抗原。结果如图8所示，表明，肝癌患者血清中存在较正常人水平高的LAPTM4B抗原。指明LAPTM4B有望成为肝癌诊断的新指标。

实施例8、以免疫共沉淀及抗体抑制分析LAPTM4B在信号转导中的作用
如实施例4制备细胞裂解液，向上清中加入针对待测靶蛋白的一抗。4℃摇动1h后，加入50 μ l Protein G-Agarose悬液，4℃摇动至少3h或过夜。
12 000g离心20秒收集免疫复合物沉淀。加入1ml洗涤缓冲液 I 重悬该复合物，4℃摇动20min，12 000g离心20秒后小心去除上清，重复该步骤一次。加洗涤缓冲液 II 重悬该复合物，4℃摇动20min，12 000g离心20秒后小心去除上清，重复该步骤一次。加洗涤缓冲液 III 重悬该复合物，4℃摇动20min，12 000g离心20秒后小心并彻底吸弃上清。沉淀中加入50 μ l 1× SDS 上样缓冲液，置100℃水浴中加热煮沸5min，使样品中的免疫复合物变性并解离。12 000g离心20秒，取上清进行SDS-PAGE电泳。BEL-7402细胞用无血清培养基在LN-1基质上分别孵育0min，10min，20min和40min。各组细胞的LAPTM4B-EC2-pAb免疫共沉淀物经Protein G-Sepharose吸附、离心分离后进行非还原性的10% SDS-PAGE，然后用p-Tyr mAb进行Western blot 分别检测LAPTM4B、FAK及MAPK的磷酸化。

BEL-7402 细胞分别与 LAPTM4B- EC2-pAb（15 μ g/ml）和 Glut2 抗体（15 μ g/ml）于 5%CO₂、37℃预温 2 h 后接种于 LN-1 基质上，孵育条件同上；以与 Glut2 抗体预温和无抗体预温的细胞作为对照组。各组细胞的蛋白质裂解液用 p-Tyr mAb 进行 Western blot 检测抗体对磷酸化的抑制作用。

当人肝癌BEL-7402细胞黏附于层粘连蛋白基质时可引起LAPTM4B-35呈峰形磷酸化，10分钟时磷酸化水平最高（图15-A），；LAPTM4B-EC2-pAb可以几近完全的抑制其磷酸化（图15-B），而无关的膜蛋白抗体Glut2则无此作用（图15-C）。LAPTM4B-24则不发生磷酸化作用。LAPTM4B-35 C端Tyr₂₈₅的磷酸化形成了与胞内信号分子SH2结构域相结合的位点；同时，LAPTM4B-35具有典型的SH3结构域的结合位点。LAPTM4B-35蛋白可能是一个在信号传导中具有关键作用的停靠蛋白或细胞质膜特定微域（microdomain）的组织者，可以募集细胞内、外相关的信号分子以完成与细胞增殖、分化及凋亡相关的信号转导。再者，人肝癌细胞在层粘连蛋白基质上黏附还可引起胞质中信号分子FAK的Tyr磷酸化（图16-A），而

LAPTM4B-EC2-pAb及针对整合素 $\alpha 6$ 胞外区的抗体与细胞预温可以阻止FAK的磷酸化, 但不影响FAK的表达水平。BEL-7402细胞在层粘连蛋白基质上黏附也可引起信号分子MAPK的Tyr磷酸化(图16-B), LAPTM4B-EC2-pAb抗体与细胞预温同样可以阻止MAPK的磷酸化, 但不影响其表达水平。这些结果指明LAPTM4B的EC2(第二胞外区)在与 $\alpha 6$ 的相互作用在启动FAK-MAPK信号转导途径中可能具有重要作用。

实施4-8的结果指明, LAPTM4B-35可以作为调节细胞增殖、分化、凋亡的药物作用的靶标。

实施例9、LAPTM4B的基因分型

设计一个以 PCR 为基础的方法对正常人和肝癌患者血液样品中的LAPTM4B进行基因分型。根据LAPTM4B基因序列3中的19bp DNA序列的侧翼序列设计合成了两个引物,

F₂: 5' GCCGACTAGGGGACTGGCGGA 3' ;

R₂: 5' CGAGAGCTCCGAGCTTCTGCC 3' 。

以基因组 DNA 为模板扩增第一外显子的部分序列。PCR 条件: 96℃预变性5分钟; 94℃ 30秒, 68℃ 30秒, 72℃ 1分钟, 35个循环; 72℃延伸5分钟。PCR产物经2%琼脂糖凝胶电泳鉴定, 结果见图10。泳道1, 6, 12, 13表示LAPTM4B*1/*1中的204bp的核苷酸片段, 泳道5, 8, 9, 14, 15表示LAPTM4B*2/*2中的223bp的核苷酸片段。泳道2, 3, 4, 7, 10, 11表示LAPTM4B*1/*2中含有204bp和223bp两个核苷酸片段。泳道M为Marker。结果表明在纯合基因对*1/*1或*2/*2中分别只有204bp或223bp DNA片段被扩增, 而在杂合基因对*1/*2中, 204bp和223bp的DNA片段同时被扩增。因此, 在中国人群中LAPTM4B的基因型可分为: LAPTM4B*1/*1, *1/*2和*2/*2(图10)。

25

实施例10、肝癌患者和正常人群LAPTM4B基因型和等位基因频率分布

本发明分析了209例中国正常人群和57例肝癌患者, 其基因型出现的频率比较见表1, 用Hardy-Weinberg方程进行数学期望分析结果。肝癌患者和正常人群的LAPTM4B等位基因*1和*2的频率有显著差异, 其比例分别为0.5175: 0.6746和0.4825: 0.3254。正常人群出现LAPTM4B等位基因*1和*2的频率分别为0.6746和0.3254, 肝癌患者出现LAPTM4B等位基因*1和*2的频率分别为0.5175和0.4825。肝癌组的基因型*1/*1($P=0.029$)和*2/*2($P=0.003$)出现的频率与其对应的正常对照组相比具有显著的统计学差异。在肝癌组中, 只有29.8%为基因型*1/*1, 正常对照组中有45.93%为基因型

*1/*1。而肝癌组中基因型*2/*2的频率为 26.32%，与对照组的 11.01%相比，其出现的频率显著增加 ($P < 0.01$)。分析表明基因型*2/*2的个体患肝癌的风险比是*2/*2型个体的 2.89 倍。因此，*LAPTM4B* *2/*2 等位基因与肝癌易感性相关。

- 5 如表 3 所示，不同基因型的患者并没有表现出肝癌级别、阶段或 HBV 感染的差异，83.3% 患者 HBV 阳性。

表 2 肝癌患者和正常人群的 *LAPTM4B* 基因型的分布

		N (%)		P 值
		对照组 B (n=209)	肝癌组 (n=57)	
LAPTM4B 基因型				
*1/*1	96 (45.93)	17 (29.82)	0.029 ^a	
*1/*2	90 (43.06)	25 (43.86)	0.914	
*2/*2	23 (11.01)	15 (26.32)	0.003 ^b	
等位基因的频率				
*1	0.6746	0.5175		
*2	0.3254	0.4825		

^a OR: 0.500, 95%CI: 0.267 - 0.939; ^b OR: 2.888, 95%CI: 1.390 - 6.003 (OR 患病风险, 和 95%CI 为置信区间)。

10 表 3 用于 *LAPTM4B* 基因分型的肝癌患者的临床资料

	<i>LAPTM4B</i> 基因型			P 值
	*1/*1	*1/*2	*2/*2	
总人数	17	25	15	
男性	14	24	12	
女性	3	1	3	NS
肿瘤的级别				
G1	0	2	0	
G2	1	4	8	
G3	7	7	4	
G4	9	12	3	NS
肿瘤的阶段				
I	0	0	0	
II	5	8	5	
III	4	7	3	
IV	8	10	7	NS

HBV 感染				
阴性	1	4	4	
阳性	13	16	10	
未确诊	3	5	1	NS

NS: 差异无显著性

实施例 11、食道癌细胞中的基因型和等位基因的频率

为了确定 *LAPTM4B* 基因型是否与其它癌症易感性相关, 对来自同一地方的 116 例正常人群和 109 例食道癌患者进行了分析。如表 4 所示, 食道癌患
5 者的 *LAPTM4B* 基因分型与其对照人群并无显著差异。表明 *LAPTM4B* 等位基因与食道癌易感性无关。

表 4 食道癌患者和正常人群的 *LAPTM4B* 基因型的分布

N (%)				P 值
对照组 B (n=209)	对照组 S (n=116)	食道癌 (n=109)		
<i>LAPTM4B</i> 基因型				
<i>*1/*1</i>	96 (45.93)	52 (44.83)	49 (44.95)	>0.05
<i>*1/*2</i>	90 (43.06)	49 (42.24)	48 (44.04)	>0.05
<i>*2/*2</i>	23 (11.01)	15 (12.93)	12 (11.01)	>0.05
等位基因的频率				
<i>*1</i>	0.6746	0.6595	0.6697	
<i>*2</i>	0.3254	0.3405	0.3303	

实施例 12、某些上皮来源肿瘤中 LAPTM-35 的表达

利用免疫组化的方法分析 LAPTM4B-35 蛋白与其它癌症的关系。具体做法
10 是, 分别取食道癌、乳腺癌、肺癌、胃癌、结肠癌和直肠癌阳性及阴性对照组织固定标本 (组织芯片), 按照如下步骤制片:

1. 标本经二甲苯脱蜡
2. 经不同浓度乙醇下行, 100%—95%—90%—80%—70%, PBS 去内
源性过氧化物酶
- 15 3. 柠檬酸钠抗原修复
4. PBS 洗两回
5. 正常山羊血清封闭
6. LAPTM4B-N₁₋₉₉pAb, 37℃保温 1 小时
7. PBS 洗 3 遍
- 20 8. HRP 标记的山羊抗兔抗体, 37℃保温 1 小时,

9. PBS 洗 3 遍
10. DAB 显色
11. 苏木金复染核
12. 经不同浓度乙醇上行脱水 (70%—80%—90%—95%—100%)
13. 封片。

5 结果如图 11 所示, 图中 A 为正常食道组织 (阴性); B 为食道癌组织 (阴性); C 为正常乳腺组织 (阴性), D 为乳腺癌组织 (阳性), E 为正常肺组织 (阴性), F 为肺癌组织 (阳性), G 为正常胃组织 (阴性), H 为胃癌组织 (阳性), 从图中可以看出, LPTM4B 在肺癌、胃癌、乳腺癌中明显表达, 而在食道癌及大肠癌中无明显表达。

10 工业应用

本发明的蛋白质可能作为癌症早期诊断的新指标, 通过采用适于在临床上广泛应用的 ELISA 法, 以及所制备的相关检测试剂盒, 可以提高某些癌症, 特别是原发性肝癌, 的早期诊断率及诊断的准确率。

15 在治疗上, 可将 *LPTM4B* 基因作为癌症治疗的靶基因: 阻遏 *LPTM4B*—35 及增强 *LPTM4B*—24 的表达可能会抑制肝癌的生长、逆转其恶性表型或延缓肝癌的发展。例如, 可通过最新发展的 siRNA 干扰技术阻抑 *LPTM4B* 基因产物 *LPTM4B*—35 的表达; 还可将 *LPTM4B*—BE cDNA 重组到工程病毒表达载体通过上调 *LPTM4B*—24 的表达进行抗肿瘤基因治疗。*LPTM4B* 蛋白也可作为药物作用的新靶点。鉴于 *LPTM4B* 蛋白可作为细胞信号转导分子的组装平台, 并具有多种信号分子的结合位点, 有可能开发出多种以 *LPTM4B* 蛋白为靶点的新药。此外, 本发明已初步证明 *LPTM4B*—EC2—pAb 抗体具有抑制肿瘤细胞增殖及阻断其信号转导的作用, 可在此基础上进一步探讨通过抗体抑制肝癌发展的可行性, 即在进一步确定效果之后研制人源化可溶性单链抗体, 用于临床患者, 并可研制肽类疫苗。如果疫苗能够研制成功不但有助于肝癌治疗, 还可用于肝癌的预防。总之, 以本发明为基础, 可能开发出一些新的抗癌途径, 作为肝癌等肿瘤治疗的重要补充, 有助于提高肝癌等的治愈率。这是一项将产生重大社会效益的工程。

30 本发明对基因组 DNA 进行了 *LPTM4B* 基因分型, 并研究了不同基因型与肝癌易感性的关系以及与其它肿瘤的易感性的关系。发现了其中一个基因型 *LPTM4B* *2/2 与肝癌的易感性密切相关, 从而为原发性肝癌易感和高危人群的筛查之准确性提供了新的更确切的标准, 对易患肝癌的高风险人群的评估和预防具有重大的意义。

权利要求书

- 1、人癌症相关基因 *LAPTM4B*, 是下列核苷酸序列之一:
 - 1) 序列表中的 SEQ ID No: 1、SEQ ID No: 2、SEQ ID No: 3、或 SEQ ID No: 6;
 - 2) 编码序列表中 SEQ ID No: 4、SEQ ID No: 5 或 SEQ ID No: 7 蛋白质序列的多核苷酸;
 - 3) 与序列表中 SEQ ID No: 1、SEQ ID No: 2、SEQ ID No: 3、或 SEQ ID No: 6 限定的 DNA 序列具有 90% 以上同源性, 且编码相同或相近的功能蛋白质的 DNA 序列。
- 2、根据权利要求 1 所述的基因, 其特征在于: 所述基因为序列表中的 SEQ ID No: 1。
- 3、根据权利要求 2 所述的基因, 其特征在于: 所述基因为序列表中的 SEQ ID No: 2。
- 4、根据权利要求 2 所述的基因, 其特征在于: 所述基因为序列表中的 SEQ ID No: 3。
- 5、根据权利要求 2 所述的基因, 其特征在于: 所述基因为序列表中的 SEQ ID No: 6。
- 6、根据权利要求 1 或 2 或 3 或 4 或 5 所述的基因, 其特征在于: 所述癌症为肝癌及某些上皮源的癌症。
- 7、人癌症相关蛋白, 是具有序列表中序列 4 或/和序列 5 或/和序列 7 氨基酸序列的蛋白质, 或者是将序列 4 或/和序列 5 或/和序列 7 的氨基酸序列经过一个或几个氨基酸残基的取代、缺失或添加且具有与序列 4 或/和序列 5 或/和序列 7 的氨基酸序列相同活性的由 4 或/和序列 5 或/和序列 7 衍生的蛋白质。
- 8、根据权利要求 7 所述的蛋白质, 其特征在于: 它是具有序列表中序列 4 氨基酸序列的蛋白质。
- 9、根据权利要求 7 所述的蛋白, 其特征在于: 它是具有序列表中序列 5 氨基酸序列的蛋白质。
- 10、根据权利要求 7 所述的蛋白, 其特征在于: 它是具有序列表中序列 7 氨基酸序列的蛋白质。
- 11、根据权利要求 7 或 8 或 9 或 10 所述的蛋白, 其特征在于: 所述癌症为肝癌及某些上皮源的癌症。
- 12、含有权利要求 1 所述基因的表达载体。

- 13、含有权利要求 1 所述基因的细胞系。
- 14、扩增权利要求 1 所述基因的引物。
- 15、人癌症相关基因 *LAPTM4B* 的启动子。
- 16、根据权利要求 15 所述的启动子，其特征在于：它具有 SEQ ID No:
5 8 的核苷酸序列。
- 17、以针对权利要求 7 所述蛋白质的各种单克隆和多克隆抗体为活性成分的试剂。
- 18、权利要求 1 所述基因在制备检测癌症的试剂中的应用。
- 19、根据权利要求 18 所述的应用，其特征在于：所述癌症为肝细胞癌及
10 某些上皮源的癌症。
- 20、权利要求 7 所述蛋白质在制备检测癌症的试剂中的应用。
- 21、根据权利要求 20 所述的应用，其特征在于：所述癌症为肝癌及某些上皮源的癌症。

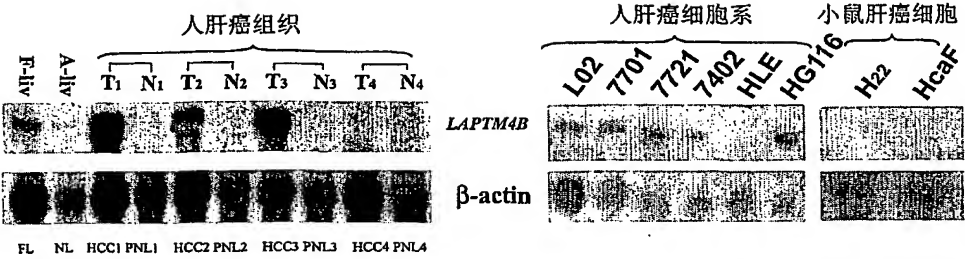


图 1-A

图 1-B

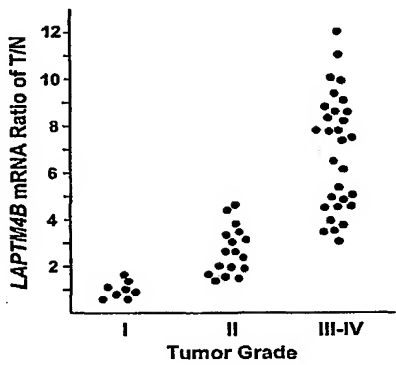


图 1-C

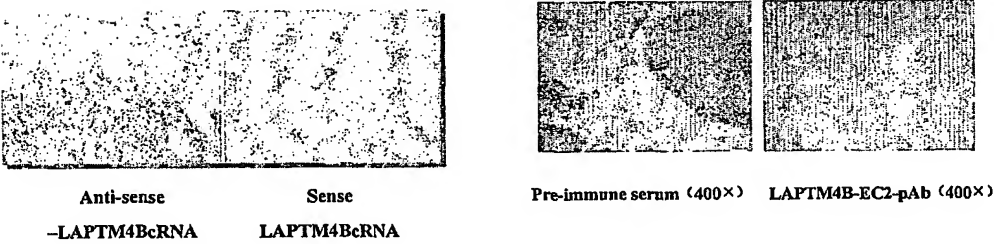


图 2-A

图 2-B



图 2-C

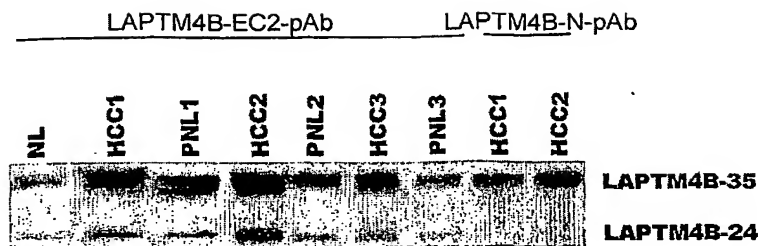


图 3

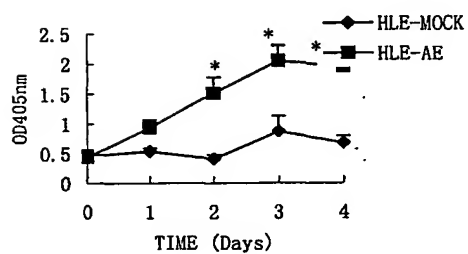


图 4

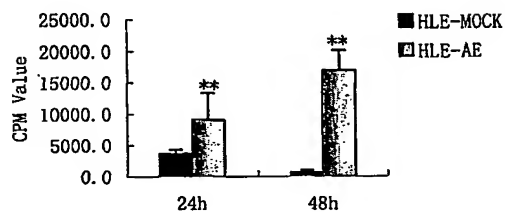


图 5



图 6

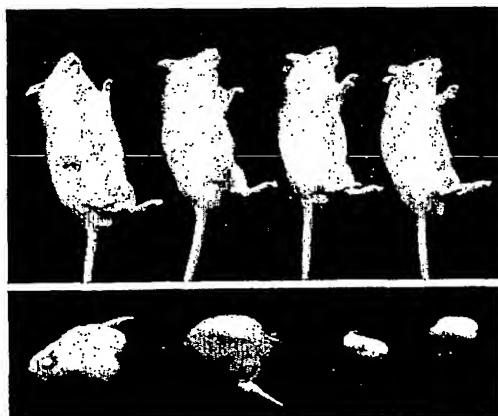


图 7

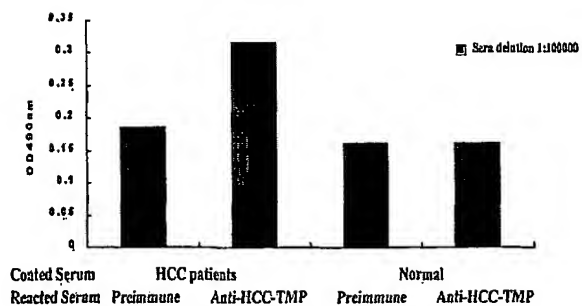


图 8

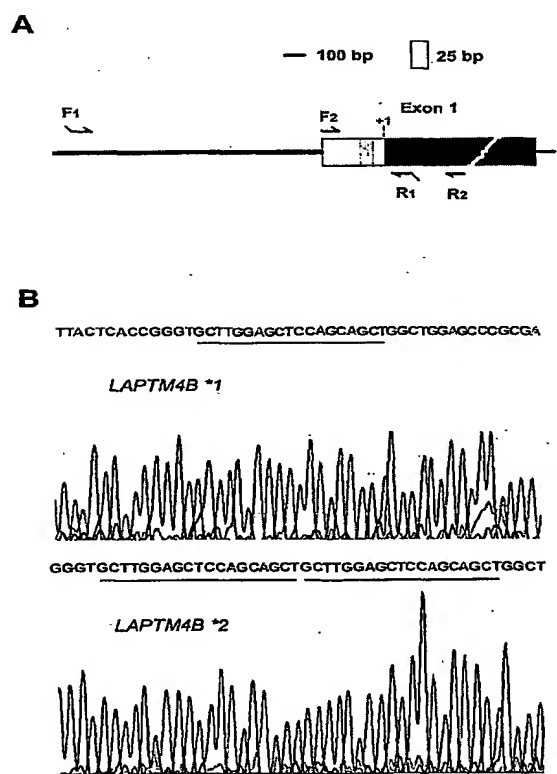


图 9

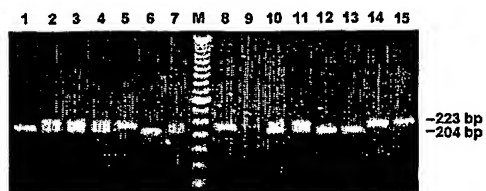


图 10

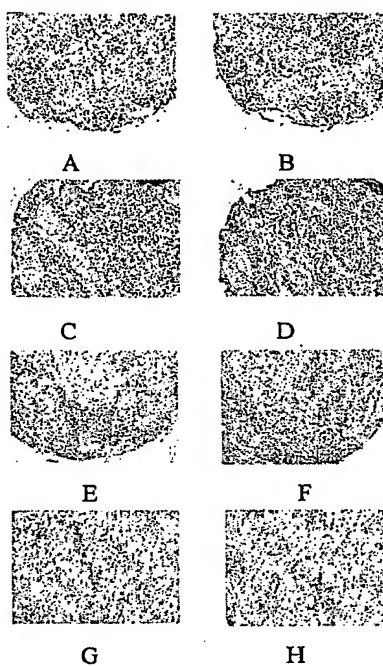
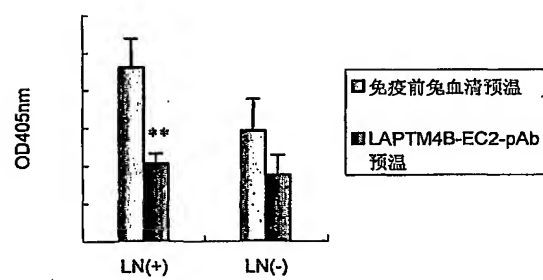


图 11



** $P < 0.01$ LAPTM4B-EC2-pAb vs. pre-immune rabbit serum

图 12

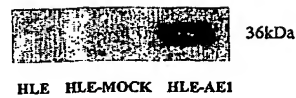


图 13-A

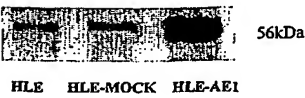


图 13-B

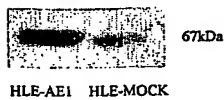


图 13-C

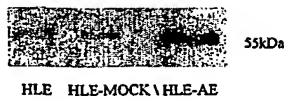


图 13-D

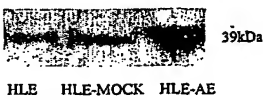


图 13-E

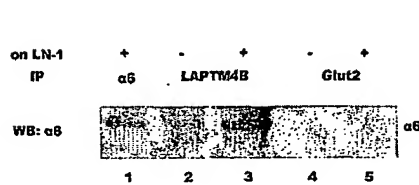


图 14-A

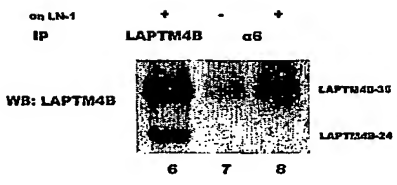


图 14-B

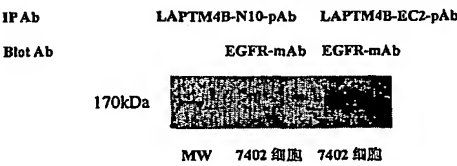


图 14-C

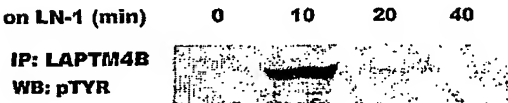


图 15-A

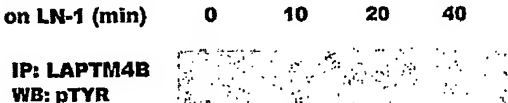


图 15-B

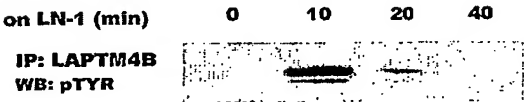


图 15-C

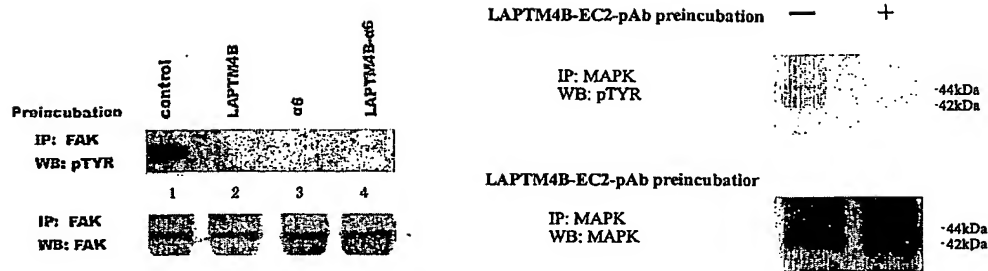


图 16-A

图 16-B

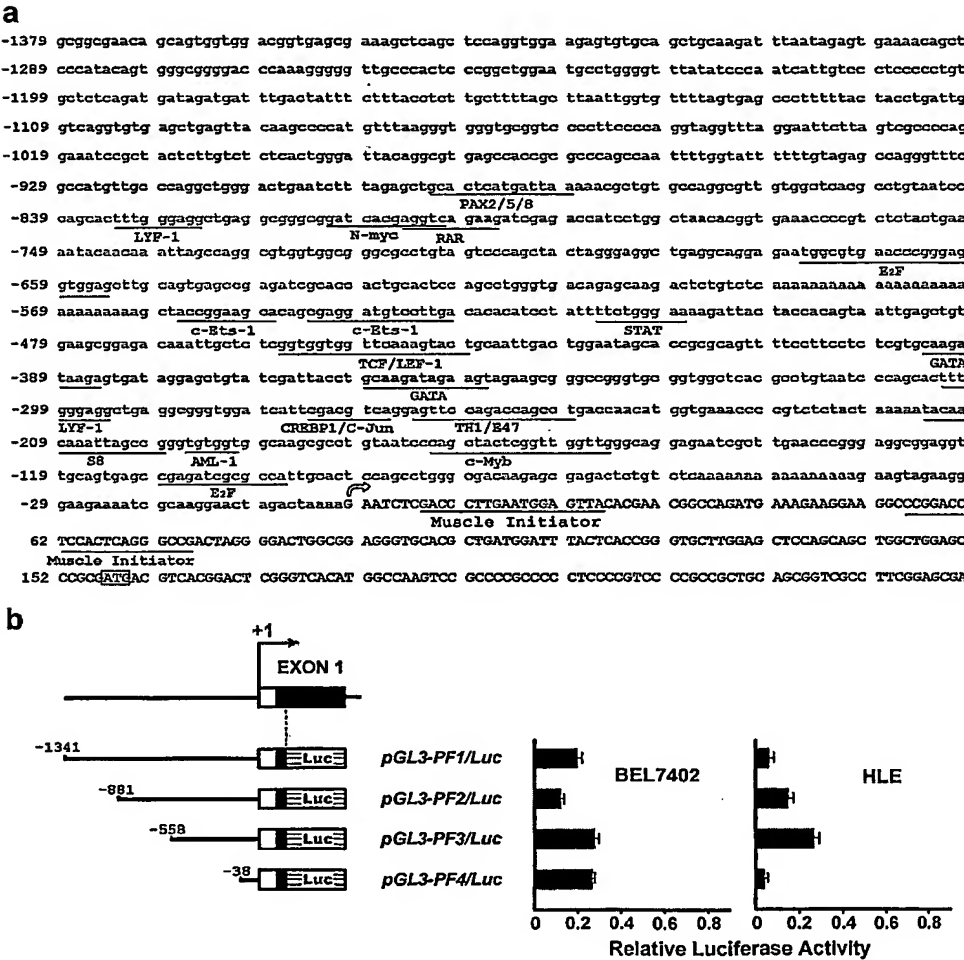


图 17

序列表

<160>8

5 <210>1

<211>954

<212>DNA

<213>人属人 (*Homo sapiens*)

10 <400>1

atgacgtcac ggactcgggt cacatggccg agtccgcccc gccccctccc cgtccccgcc	60
gctgcagccg tcgccttcgg agcgaagggt accgacccgg cagaagctcg gagctctcgg	120
ggtatcgagg aggcaggccc gcgggcgcac gggcgagcgg gccgggagcc ggagcggcgg	180
aggagccggc agcagcggcg cggcgggctc caggcgaggc ggtcgacgct cctgaaaact	240
15 tgcgcgcgcg ctgcgcgccac tgcgcccga gcgatgaaga tggcgcgcc ctggacgcgg	300
ttctactcca acagctgctg cttgtgctgc catgtccgca ccggcaccat cctgctcggc	360
gtctggtatc tgatcatcaa tgctgtgga ctgttgattt tattgagtgc cctggctgat	420
ccgatcagat ataactttc aagttctgaa ctgggaggtg actttgagtt catggatgat	480
gccaacatgt gcattgccat tgcgatttct cttctcatga tcctgatatg tgctatggct	540
20 acttacggag cgtacaagca acgcgcagcc tggatcatcc cattcttctg ttaccagatc	600
tttgactttg ccctgaacat gttggttgca atcactgtgc ttatttatcc aaactccatt	660
caggaataca tacggcaact gcctcctaata tttccctaca gagatgatgt catgtcagtg	720
aatcctacct gtttggtcct tattattctt ctgtttatta gcattatctt gacttttaag	780
ggttacttga ttagctgtgt ttggaactgc taccgataca tcaatggtag gaactcctct	840
25 gatgtcctgg tttatgttac cagcaatgac actacggtgc tgctaccccc gtatgatgat	900
gccactgtga atggtgctgc caaggagcca ccgccacctt acgtgtctgc ctaa	954

<210>2

<211>1440

30 <212>DNA

<213>人属人 (*Homo sapiens*)

<400>2

gccgactagg ggactggcgg aggggtgcacg ctgatggatt tactcaccgg gtgcttgag	60
35 ctccagcagc tggtcggagc ccgcgatgac gtcacggact cgggtcacat ggccgagtcc	120
gccccgcccc ctccccgtcc ccgccgtgc agccgtcgcc ttcggagcga aggggtaccga	180

cccggcagaa gctcggagct ctcggggtat cgaggaggca ggcccgcggg cgcacgggcg 240
 agcgggccgg gagccggagc ggcggaggag ccggcagcag cggcgcggcg ggctccaggc 300
 gaggcggctg acgctcctga aaacttgccg gcgcgctcgc gccactgcgc ccggagcgat 360
 gaagatggtc gcgccctgga cgcggttcta ctccaacagc tgctgcttgt gctgccatgt 420
 5 ccgcaccggc accatcctgc tcggcgtctg gtatctgac atcaatgctg tggtagtgtt 480
 gattttattg agtgccctgg ctgatccga tcagtataac ttttcaagtt ctgaactggg 540
 aggtgacttt gagttcatgg atgatccaa catgtgcatt gccattgcga tttctcttct 600
 catgacctg atatgtgcta tggctactta cggagcgtac aagcaacgcg cagcctggat 660
 catccattc ttctgttacc agatctttga ctttgccctg aacatgttgg ttgcaatcac 720
 10 tgtgcttatt tatccaaact ccattcagga atacatacgg caactgcctc ctaattttcc 780
 ctacagagat gatgtcatgt cagtgaatcc tacctgtttg gtccttatta ttcttctgtt 840
 tattagcatt atcttgactt ttaagggtta cttgattagc tgtgtttgga actgctaccg 900
 atacatcaat ggtaggaact cctctgatgt cctggtttat gttaccagca atgacactac 960
 ggtgctgcta ccccgctatg atgatgccac tgtgaatggt gctgccaagg agccaccgcc 1020
 15 acctacgtg tctgcctaag ccttcaagtg ggcggagctg agggcagcag cttgactttg 1080
 cagacatctg agcaatagtt ctgttatttc acttttgcca tgagcctctc tgagcttgtt 1140
 tgttgctgaa atgctacttt ttaaaattta gatgttagat tgaaaactgt agttttcaac 1200
 atatgctttg ctggaacact gtgatagatt aactgtagaa ttcttctgt acgattgggg 1260
 atataatggg cttcactaac cttccctagg cattgaaact tccccaaat ctgatggacc 1320
 20 tagaagtctg cttttgtacc tgctgggccc caaagtggg catttttctc tctgttcctt 1380
 ctcttttgaa aatgtaaaat aaaacaaaa atagacaaa aaaaaaaaa aaaaaaaaa 1440

<210>3

<211>2169

25 <212>DNA

<213>人属人 (*Homo sapiens*)

<400>3

gccgactagg ggactggcgg aggggtgcacg ctgatggatt tactcaccgg gtgcttggag 60
 30 ctccagcagc tggctggagc ccgcgatgac gtcacggact cgggtcacat ggccgagtcc 120
 gccccgcccc ctccccgtcc ccgccgtgc agccgtcgc ttcggagcga agggtagcga 180
 cccggcagaa gctcggagct ctcggggtat cgaggaggca ggcccgcggg cgcacgggcg 240
 agcgggccgg gagccggagc ggcggaggag ccggcagcag cggcgcggcg ggctccaggc 300
 gaggcggctg acgctcctga aaacttgccg gcgcgctcgc gccactgcgc ccggagcgat 360
 35 gaagatggtc gcgccctgga cgcggttcta ctccaacagc tgctgcttgt gctgccatgt 420
 ccgcaccggc accatcctgc tcggcgtctg gtatctgac atcaatgctg tggtagtgtt 480

gattttattg agtgccctgg ctgatccgga tcagtataac ttttcaagtt ctgaactggg 540
 aggtgacttt gagttcatgg atgatgccaa catgtgcatt gccattgcga tttctcttct 600
 catgatcctg atatgtgcta tggctactta cggagcgtac aagcaacgcg cagcctggat 660
 catcccatc tttctgtacc agatctttga ctttgccctg aacatgttgg ttgcaatcac 720
 5 tgtgcttatt tatccaaact ccattcagga atacatacgg caactgcctc ctaattttcc 780
 ctacagagat gatgtcatgt cagtgaatcc tacctgtttg gtccttatta tttctctgtt 840
 tattagcatt atcttgactt ttaagggta cttgattagc tgtgtttgga actgctaccg 900
 atacatcaat ggtaggaact cctctgatgt cctggtttat gttaccagca atgacactac 960
 ggtgctgcta cccccgtatg atgatgccac tgtgaatggt gctgccaaagg agccaccgcc 1020
 10 accttacgtg tctgcctaag cttcaagtg ggaggagctg agggcagcag cttgactttg 1080
 cagacatctg agcaatagtt ctgttatttc acttttgcca tgagcctctc tgagcttggt 1140
 tgttgctgaa atgctacttt ttaaaattta gatgttagat tgaaaactgt agttttcaac 1200
 atatgctttg ctggaacact gtgatagatt aactgtagaa ttcttcctgt acgattgggg 1260
 atataatggg cttcactaac cttccctagg cattgaaact tcccccaat ctgatggacc 1320
 15 tagaagtctg cttttgtacc tgctgggccc caaagttggg catttttctc tctgttcct 1380
 ctcttttgaa aatgtaaaat aaaacaaaa atagacaact tttcttcag ccattccagc 1440
 atagagaaca aaaccttatg gaaacaggaa tgtcaattgt gtaatcattg ttctaattag 1500
 gtaaatagaa gtccttatgt atgtgttaca agaatttccc ccacaacatc ctttatgact 1560
 gaagttcaat gacagtttgt gtttggtggt aaaggatttt ctccatggcc tgaattaaga 1620
 20 ccattagaaa gcaccaggcc gtgggagcag tgaccatctg ctgactgttc ttgtggatct 1680
 tgtgtccagg gacatggggg gacatgcctc gtatgtgtta gaggggtggaa tggatgtgtt 1740
 tggcgctgca tgggatctgg tgccctctt ctcttgatt cacatccca cccagggcc 1800
 gcttttacta agtgttctgc cctagattgg ttcaaggagg tcatccaact gactttatcg 1860
 agtgggaattg ggatatattt gatatacttc tgcctaaca catggaaaag ggttttcttt 1920
 25 tccctgcaag ctacatccta ctgctttgaa cttccaagta tgtctagtca ctttttaaaa 1980
 tgtaaacatt ttcagaaaaa tgaggattgc cttccttgta tgcgcttttt accttgacta 2040
 cctgaattgc aagggaattt tatatattca tatgttaca agtcagcaac tctcctgttg 2100
 gttcattatt gaatgtgctg taaattaagt tgtttgaat taaaacaagg tttgccaca 2160
 aaaaaaaaa 2169

30

<210>4

<211>317

<212>PRT

<213>人属人 (*Homo sapiens*)

35

<400>4

	Met	Thr	Ser	Arg	Thr	Arg	Val	Thr	Trp	Pro	Ser	Pro	Pro	Arg	Pro
	1				5					10					15
	Leu	Pro	Val	Pro	Ala	Ala	Ala	Ala	Val	Ala	Phe	Gly	Ala	Lys	Gly
					20					25					30
5	Thr	Asp	Pro	Ala	Glu	Ala	Arg	Ser	Ser	Arg	Gly	Ile	Glu	Glu	Ala
					35					40					45
	Gly	Pro	Arg	Ala	His	Gly	Arg	Ala	Gly	Arg	Glu	Pro	Glu	Arg	Arg
					50					55					60
	Arg	Ser	Arg	Gln	Gln	Arg	Arg	Gly	Gly	Leu	Gln	Ala	Arg	Arg	Ser
10					65					70					75
	Thr	Leu	Leu	Lys	Thr	Cys	Ala	Arg	Ala	Arg	Ala	Thr	Ala	Pro	Gly
					80					85					90
	Ala	Met	Lys	Met	Val	Ala	Pro	Trp	Thr	Arg	Phe	Tyr	Ser	Asn	Ser
					95					100					105
15	Cys	Cys	Leu	Cys	Cys	His	Val	Arg	Thr	Gly	Thr	Ile	Leu	Leu	Gly
					110					115					120
	Val	Trp	Tyr	Leu	Ile	Ile	Asn	Ala	Val	Val	Leu	Leu	Ile	Leu	Leu
					125					130					135
	Ser	Ala	Leu	Ala	Asp	Pro	Asp	Gln	Tyr	Asn	Phe	Ser	Ser	Ser	Glu
20					140					145					150
	Leu	Gly	Gly	Asp	Phe	Glu	Phe	Met	Asp	Asp	Ala	Asn	Met	Cys	Ile
					155					160					165
	Ala	Ile	Ala	Ile	Ser	Leu	Leu	Met	Ile	Leu	Ile	Cys	Ala	Met	Ala
					170					175					180
25	Thr	Tyr	Gly	Ala	Tyr	Lys	Gln	Arg	Ala	Ala	Trp	Ile	Ile	Pro	Phe
					185					190					195
	Phe	Cys	Tyr	Gln	Ile	Phe	Asp	Phe	Ala	Leu	Asn	Met	Leu	Val	Ala
					200					205					210
	Ile	Thr	Val	Leu	Ile	Tyr	Pro	Asn	Ser	Ile	Gln	Glu	Tyr	Ile	Arg
30					215					220					225
	Gln	Leu	Pro	Pro	Asn	Phe	Pro	Tyr	Arg	Asp	Asp	Val	Met	Ser	Val
					230					235					240
	Asn	Pro	Thr	Cys	Leu	Val	Leu	Ile	Ile	Leu	Leu	Phe	Ile	Ser	Ile
					245					250					255
35	Ile	Leu	Thr	Phe	Lys	Gly	Tyr	Leu	Ile	Ser	Cys	Val	Trp	Asn	Cys
					260					265					270

Tyr Arg Tyr Ile Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr
 275 280 285
 Val Thr Ser Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp
 290 295 300
 5 Ala Thr Val Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val
 305 310 315
 Ser Ala
 317
 10 <210>5
 <211>226
 <212>PRT
 <213>人属人 (*Homo sapiens*)
 15 <400>5
 Met Lys Met Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys
 1 5 10 15
 Cys Leu Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val
 20 20 25 30
 Trp Tyr Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser
 35 40 45
 Ala Leu Ala Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu
 50 55 60
 Gly Gly Asp Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala
 25 65 70 75
 Ile Ala Ile Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr
 80 85 90
 Tyr Gly Ala Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe
 95 100 105
 30 Cys Tyr Gln Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile
 110 115 120
 Thr Val Leu Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln
 125 130 135
 Leu Pro Pro Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn
 35 140 145 150
 Pro Thr Cys Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile

	155	160	165
	Leu Thr Phe Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr		
	170	175	180
	Arg Tyr Ile Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val		
5	185	190	195
	Thr Ser Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala		
	200	205	210
	Thr Val Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser		
	215	220	225

10 Ala
226

<210>6

<211>2264

15 <212>DNA

<213>人属人 (*Homo sapiens*)

<400>6

	gaatctcgac ccttgaatgg agttacacga acggccagat gaaagaagga aggcccggac	60
20	ctccactcag ggccgactag gggactggcg gaggggtgcac gctgatggat ttactcaccg	120
	ggtgcttggg gctccagcag ctgcttggag ctccagcagc tggctggagc ccgcgatgac	180
	gtcacggact cgggtcacat ggccgagtcc gccccgcccc ctccccgtcc ccgcccgtgc	240
	agccgtcgcc ttcggagcga aggggtaccga cccggcagaa gctcggagct ctcggggtat	300
	cgaggaggca ggcccgcggg cgcacgggagc agcggggcgg gagccggagc ggccggaggag	360
25	ccggcagcag cggcgcgggc ggctccaggc gaggcggtcg acgctcctga aaacttgcgc	420
	gcgcgctcgc gccactgcgc ccggagcgat gaagatggtc gcgccctgga cgcggttcta	480
	ctccaacagc tgcgtcttgt gctgccatgt ccgcaccggc accatcctgc tcggcgtctg	540
	gtatctgata atcaatgctg tggactgtt gattttattg agtgccctgg ctgatccgga	600
	tcagtataac ttttcaagtt ctgaactggg aggtgacttt gagttcatgg atgatgcaa	660
30	catgtgcatt gccattgcga tttctcttct catgatcctg atatgtgcta tggctactta	720
	cggagcgtac aagcaacgcg cagcctggat catccattc ttctgttacc agatctttga	780
	ctttgccctg aacatgttgg ttgcaatcac tgtgcttatt tatccaaact ccattcagga	840
	atacatacgg caactgcctc ctaattttcc ctacagagat gatgtcatgt cagtgaatcc	900
	tacctgtttg gtccttatta ttcttctgtt tattagcatt atcttgactt ttaagggtta	960
35	cttgattagc tgtgtttgga actgctaccg atacatcaat ggtaggaact cctctgatgt	1020
	cctggtttat gttaccagca atgacactac ggtgctgcta cccccgtatg atgatgccac	1080

tgtgaatggt gctgccaaagg agccaccgcc accttacgtg tctgcctaag ccttcaagtg 1140
 ggcggagctg agggcagcag cttgactttg cagacatctg agcaatagtt ctgttatttc 1200
 acttttgcca tgagcctctc tgagcttggt tgttgctgaa atgctacttt ttaaaattta 1260
 gatgttagat tgaaaactgt agttttcaac atatgctttg ctggaacact gtgatagatt 1320
 5 aactgtagaa ttcttcctgt acgattgggg atataatggg cttcactaac cttccctagg 1380
 cattgaaact tccccaaat ctgatggacc tagaagtctg cttttgtacc tgctggggcc 1440
 caaagttggg ctttttctc tctgttcctt ctcttttgaa aatgtaaaat aaaacaaaaa 1500
 atagacaact ttttcttcag ccattccagc atagagaaca aaaccttatg gaaacaggaa 1560
 tgtcaattgt gtaatcattg ttctaattag gtaaatagaa gtccttatgt atgtgttaca 1620
 10 agaatttccc ccacaacatc ctttatgact gaagttcaat gacagtgtgt gtttggtggt 1680
 aaaggatttt ctccatggcc tgaattaaga ccattagaaa gcaccaggcc gtgggagcag 1740
 tgaccatctg ctgactgttc ttgtggatct tgtgtccagg gacatggggg gacatgcctc 1800
 gtatgtgtta gagggtggaa tggatgtgtt tggcgtgca tgggatctgg tgcccctctt 1860
 ctcttggtt caccatcccca cccaggggcc gcttttacta agtgttctgc cctagattgg 1920
 15 ttcaaggagg tcatccaact gactttatcg agtggattg ggatataatt gatatacttc 1980
 tgcctaacaa catggaaaag ggttttcttt tcctgcaag ctacatccta ctgctttgaa 2040
 cttccaagta tgtctagtca ccttttaaaa tgtaaacatt ttcagaaaaa tgaggattgc 2100
 cttccttgta tgcgcttttt accttgacta cctgaattgc aagggtttt tatatatcca 2160
 tatgttacia agtcagcaac tctcctgttg gttcattatt gaatgtgctg taaattaagt 2220
 20 tgtttgcaat taaaacaagg tttgccaca aaaaaaaaaa aaaa 2264

<210>7

<211>370

<212>PRT

25 <213>人属人 (*Homo sapiens*)

<400>7

Met Glu Leu His Glu Arg Pro Asp Glu Arg Arg Lys Ala Arg Thr
 1 5 10 15
 30 Ser Thr Gln Gly Arg Leu Gly Asp Trp Arg Arg Val His Ala Asp
 20 25 30
 Gly Phe Thr His Arg Val Leu Gly Ala Pro Ala Ala Ala Trp Ser
 35 40 45
 Ser Ser Ser Trp Leu Glu Pro Ala Met Thr Ser Arg Thr Arg Val
 35 50 55 60
 Thr Trp Pro Ser Pro Pro Arg Pro Leu Pro Val Pro Ala Ala Ala

	65	70	75
	Ala Val Ala Phe Gly Ala Lys Gly Thr Asp Pro Ala Glu Ala Arg		
	80	85	90
	Ser Ser Arg Gly Ile Glu Glu Ala Gly Pro Arg Ala His Gly Arg		
5	95	100	105
	Ala Gly Arg Glu Pro Glu Arg Arg Arg Ser Arg Gln Gln Arg Arg		
	110	115	120
	Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu Lys Thr Cys Ala		
	125	130	135
10	Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys Met Val Ala Pro		
	140	145	150
	Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu Cys Cys His Val		
	155	160	165
	Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr Leu Ile Ile Asn		
15	170	175	180
	Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu Ala Asp Pro Asp		
	185	190	195
	Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly Asp Phe Glu Phe		
	200	205	210
20	Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala Ile Ser Leu Leu		
	215	220	225
	Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly Ala Tyr Lys Gln		
	230	235	240
	Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr Gln Ile Phe Asp		
25	245	250	255
	Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val Leu Ile Tyr Pro		
	260	265	270
	Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro Pro Asn Phe Pro		
	275	280	285
30	Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr Cys Leu Val Leu		
	290	295	300
	Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr Phe Lys Gly Tyr		
	305	310	315
	Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr Ile Asn Gly Arg		
35	320	325	330
	Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser Asn Asp Thr Thr		

	335	340	345
	Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val Asn Gly Ala Ala		
	350	355	360
	Lys Glu Pro Pro Pro Tyr Val Ser Ala		
5	365	370	

<210>8

<211>1341

<212>DNA

10 <213>人属人 (*Homo sapiens*)

<400>8

	gctccaggtg gaagagtgtg cagctgcaag atttaataga gtgaaaacag ctcccataca	60
	gtggcgggg acccaaagg ggttgccac tccggctgg aatgcctgg gtttatatcc	120
15	caatcattgt ccctccccct gtgctctcag atgatagatg atttgactat ttctttacct	180
	cttgctttta gcttaattgg tgttttagtg agccctttt actacctgat tggtcaggtg	240
	tgagctgagt tacaagcccc atgtttaagg gtgggtgcgg tccccctccc caggtaggtt	300
	taggaattct tagtcgcccc aggaaatccg ctactcttgt ctctcactgg gattacaggc	360
	gtgagccacc gcgcccagcc aatttttgta tttttttag agccagggtt tcgccatggt	420
20	gccaggtg ggactgaatc tttagagctg cactcatgat taaaaacgt gtgccaggcg	480
	tttggtgctc cgcctgtaat ccagcactt tgggaggctg aggcgggcgg atcacagggt	540
	cagaagatcg agaccatcct ggctaacacg gtgaaacccc gtctctactg aaaatacaac	600
	aaattagcca ggcgtggtgg cgggcgcctg tagtcccagc tactaggag gctgaggcag	660
	gagaatggcg tgaaccggg aggtggagct tgcagtgagc cgagatcgca cactgcact	720
25	ccagcctggg tgacagagca agactctgtc tcaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	780
	agctaccgga agcacagcga ggatgtcctt gacacacatc ctattttctg ggaaaagatt	840
	actaccacag taattgagct gtgaagcgga gacaaattgc tctcgtggt ggttcaaagt	900
	actgcaattg actggaatag caccgcgcag ttttccttcc tctcgtgcaa gataagagt	960
	ataggagctg tatcgattac ctgcaagata gaagtagaag cgggccgggt gcggtggctc	1020
30	acgcctgtaa tcccagcact ttgggaggct gaggcgggtg gatcattcga cgtcaggagt	1080
	tccagaccag cctgaccaac atggtgaaac cccgtctcta ctaaaaatac aacaaattag	1140
	cgggtgtgg tggcaagcgc ctgtaatccc agctactcgg ttggttgggc aggagaatcg	1200
	cttgaacccg ggaggcggag gttgcagtga gccgagatcg cgccattgca ctccagcctg	1260
	ggcgacaaga gcgagactct gtctcaaaaa aaaaaaaaaa agaagtagaa gggaagaaaa	1320
35	tcgcaaggaa ctgactaaa a	1341

国际检索报告

国际申请号

PCT/CN 03/01109

A. 主题的分类

Int Cl⁷: C12N 15/12, C07K 14/435, C12N 15/63, 15/11, C12Q 1/68, G01N 33/574, 33/577, 33/68

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

Int Cl⁷: C12N, C07K, C12Q, G01N

包含在检索领域中的除最低限度文献以外的检索文献

Int Cl⁷: C07H, C12P, A61K, A61P

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

NCBI、EPDOC、WPI、PAJ、CPRS、CNKI、CA、MEDLINE、BA

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求编号
X	Genebank 登录号: AC098481、AC107048、AC108164、AC114876、AF317417、AJ276485、AK075326、AL136942、AL591704、AP002906、AP003357	1
A	Molecular Carcinogenesis, 第 27 卷第 3 期, 2000 年 3 月 31 日 (31.03.00), Recio J A 等, "The human PCPH proto-oncogene: cDNA identification, primary structure, chromosomal mapping, and expression in normal and tumor cells", 第 229-236 页	1-21
A	WO A1 9109045 (US DEPT OF COMMERCE、US NAT CANCER INST、NAT INST OF HEALTH、US SEC OF COMMERCE、US DEPT HEALTH & HUMAN SERVICES) 1991 年 06 月 27 日 (27.06.91) 全文	1-21

☐ 其余文件在 C 栏的续页中列出。☒ 见同族专利附件。

* 引用文件的专用类型:

“A” 明确叙述了被认为不是特别相关的一般现有技术的文件

“E” 在国际申请日的当天或之后公布的在先的申请或专利

“L” 可能引起对优先权要求的怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理

“X” 特别相关的文件, 仅仅考虑该文件, 权利要求所记载的发明就不能认为是新颖的或不能认为是有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 权利要求记载的发明不具有创造性

“&” 同族专利成员的文件

国际检索实际完成的日期

22.03 月 2004 (22.03.04)

国际检索报告邮寄日期

01 · 4月 2004 (01 · 04 · 2004)

国际检索单位名称和邮寄地址

ISA/CN

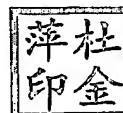
中国北京市海淀区西土城路 6 号(100088)

传真号: 86-10-62019451

受权官员

杜金萍

电话号码: 86-10-62085297



国际检索报告
关于同族专利成员的情报

国际申请号
PCT/CN 03/01109

检索报告中引用的 专利文件	公布日期	同族专利成员	公布日期
WO 9109045 A1	27.06 月 1991 (27.06.91)	CA 2071924 A	20.06 月 1991 (20.06.91)
		AU 6975291 A	18.07 月 1991 (18.07.91)
		CN 1054615 A	18.09 月 1991 (18.09.91)
		EP 0506766 A	07.10 月 1992 (07.10.92)
		JP 4506459 T	12.11 月 1992 (12.11.92)
		AU 636435 B	29.04 月 1993 (29.04.93)
		US 5403926 A	04.04 月 1995 (04.04.95)
		KR 9609744 B	24.07 月 1996 (24.07.96)
		US 5702907 A	30.12 月 1997 (30.12.97)
		US 5811262 A	22.09 月 1998 (22.09.98)
		AT 180489 T	15.06 月 1999 (15.06.99)
		DE 69033128 D	01.07 月 1999 (01.07.99)
		JP 3091870 B2	25.09 月 2000 (25.09.00)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record.**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☒ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.